



# Gradient-Based Tone Mapping for Rate-Distortion Optimized Backward-Compatible High Dynamic Range Compression

David Gommelet, Aline Roumy, Christine Guillemot, Michaël Ropert, Julien Le Tanou

## ► To cite this version:

David Gommelet, Aline Roumy, Christine Guillemot, Michaël Ropert, Julien Le Tanou. Gradient-Based Tone Mapping for Rate-Distortion Optimized Backward-Compatible High Dynamic Range Compression. IEEE Transactions on Image Processing, 2017, 10.1109/TIP.2017.2740159. hal-01590258

**HAL Id: hal-01590258**

**<https://inria.hal.science/hal-01590258>**

Submitted on 19 Sep 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Gradient-Based Tone Mapping for Rate-Distortion Optimized Backward-Compatible High Dynamic Range Compression

David Gommelet, Aline Roumy, *Member, IEEE*, Christine Guillemot, *Fellow, IEEE*, Michaël Ropert, and Julien Le Tanou,

**Abstract**—This paper addresses the problem of designing a global tone mapping operator for rate distortion optimized backward compatible compression of HDR images. We address the problem of tone mapping design for two different use cases leading to two different minimization problems. The first problem considered is the minimization of the distortion on the reconstructed HDR signal under a rate constraint on the SDR layer. The second problem remains the same minimization with an additional constraint to preserve a good quality for the SDR signal. Both the distortion and the rate are expressed as a function of the spatial gradient in the HDR images. Experiments show that the proposed rate and distortion models based on the HDR image gradient accurately predict the real image rate and distortion measures. Experimental results show that for the first minimization, the optimal rate-distortion performances are achieved, and that the second optimization yields the best trade-off between rate-distortion performance and quality preservation of the SDR signal.

**Index Terms**—High dynamic range (HDR), Compression, Tone-Mapping (TMO), Backward-Compatible, HEVC, Intra-Frame R-D model, Gradient

## I. INTRODUCTION

WITH pupil's adaptation, the human eye can see 14 orders of luminance magnitude (going from  $10^{-6}$  to  $10^8$   $cd/m^2$  or nits [1]) and without pupil's adaptation the human eye can instantaneously see around 5 orders of luminance magnitude [2]. While Standard Dynamic Range (SDR) displays only reproduce 3 orders of magnitude, ranging from 0.1 to 150  $cd/m^2$ , High Dynamic Range (HDR) aims at capturing and reproducing all visible luminance values. To do so, the real<sup>1</sup> physical measures are first sensed, leading to a 32 bit representation per RGB component (See Fig-1). Then, for efficient storage and manipulation, several formats have been proposed such as RGBE [3] and OpenEXR [4] formats. They consist of first reducing the range, only for OpenEXR ( $65.10^3$   $cd/m^2$ ), by applying a linear map and then reducing the storage by applying a quantizer. This leads to a 32 bit per pixel representation for RGBE (8 bits for each RGB component and 8 bits for a shared exponent) and a 48 bit per pixel representation for OpenEXR (16 bits for each

component). In the following, we refer to this representation as *Linear Light (LL) HDR* (See Fig-1).

Legacy video codecs are not well adapted to these LL HDR content. Indeed, they process integer values and not half-floating point representations. Moreover, legacy video codecs usually optimize the compression based on a mean square error (MSE) criterion. Therefore, the codec needs to be supplied with uniformly quantifiable data such that a one bit quantization error has the same perceptual effect in the bright and dark areas. This is done through Perceptual Quantizer Transfer Functions (PQTFs). Several solutions exist like [5] where the logluv PQTF [6] is used to convert the OpenEXR files into 14 bits for the luma representation. In the MPEG call for evidence (CFE) [7], the SMPTE-2084 [8] is used to transform the LL HDR content into 10 bits before the encoding process using HEVC codec. As a response to the CFE, many other PQTFs were proposed among which the ARIB STD-67 (HLG) [9], [10] has also been retained and is studied in the HEVC standard [11]. In Fig-1, we refer to this representation as PQTF-HDR. However, in the remainder of this paper, we will only consider PQTF-HDR. Therefore, without loss of clarity, we refer now to PQTF-HDR as HDR content. Note that, in general, quantizing an image before encoding decreases the rate-distortion performances [12]. However, it is shown in [8] that, in the case where the image is quantized with at least 12 bits, this prior quantization has no perceptual impact.

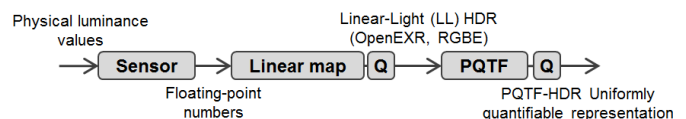


Fig. 1. HDR representation scheme

This HDR content can then be compressed with a legacy video codec [13], [14]. However, one should not only consider the problem of efficiently compressing HDR content, but also tackle the problem of backward compatibility. Indeed, most consumers are equipped with SDR displays, that cannot handle HDR content, and they will slowly renew their display to an HDR one. It is therefore of great importance to deliver an HDR signal that can be decoded by both SDR and HDR displays. A first solution is to send both SDR and HDR representations (simulcast), but this is known to be suboptimal in terms of transmission cost [15]. Instead, a scalable video codec encodes

D. Gommelet, A. Roumy and C. Guillemot are with Inria, Rennes-Bretagne-Atlantique, Rennes 35042, France (e-mail: firstname.lastname@inria.fr).

D. Gommelet, M. Ropert and J.Le Tanou are with Ericsson, Saint-Jacques-de-la-Lande 35136, France (e-mail: firstname.lastname@ericsson.com).

Manuscript received ...; revised ...

<sup>1</sup>real in the mathematical meaning

an HDR content into a single bitstream that can be decoded by both SDR and HDR displays. In a scalable scheme, the HDR input is first tone mapped with a global tone mapping operator (TMO) into an 8 bit SDR version which is encoded with a legacy codec such as HEVC or H.264 and sent along with metadata [11], [16]. Then, the decoded SDR content is inverse tone mapped to yield the HDR reconstructed content. One can also send an enhancement layer containing the HDR residue to improve the reconstruction.

Thus, the TMO is an essential component for backward compatibility and can be designed for different purposes. First, it can be designed to preserve the HDR artistic intent in the tone mapped SDR content. Many TMOs have been developed for this purpose. Some examples that consistently perform well [17], [18] are the photographic tone reproducer (PTR) [19] and the display adaptive tone mapping (DATM) [20]. Secondly, in the context of compression, the tone mapping can be designed to optimize the compression performances [21], [22], [23]. In [21] a TMO is designed to minimize the distortion of the decoded HDR content. However, this distortion model assumes that (i) the quantization error is independent of the signal and that (ii) the variance of this quantization error does not depend on the statistics of the signal. Instead, we propose a new distortion model based on the image gradient and that better reflects the rate-distortion performance of actual encoders compared with state of the art models. Closed form expressions are derived for this model which relies on the statistics of the HDR image and on the applied TMO. Since the approach in [21] does not consider the rate cost of the SDR content, the designed TMO may lead to SDR data with increased spatial complexity, hence with high rate cost. The authors in [22] and [23] cope with this limitation by minimizing the HDR distortion under the constraint of the SDR data rate. In [22], the SDR rate is modeled as a function of the total variation of the SDR signal and in [23] the SDR rate is modeled as a function of the SDR entropy.

In this paper, we also address this minimization problem. However, we propose a new rate model based on the image gradient. The gradient model is shown to provide the highest correlation with the actual SDR rate. These two new gradient-based models hold when an image is first transformed with an invertible piecewise affine function and then compressed with a predictive encoder such as the intra mode of HEVC.

Optimizing the TMO for compression purposes as in [21], where the authors minimize only the distortion on the inverse tone mapped HDR, may yield SDR content which does not preserve the artistic intent of the HDR. Thus the approach of [21] is extended in [24], [25], [26], [27] by adding a new constraint on the SDR perceptual quality. It consists of a MSE between the tone mapped signal and a reference SDR version. However this constraint is computed with a coarse assumption. Instead, we propose a new model for the SDR quality constraint using a weaker assumption. As opposed to the proposed solution, the influence of the rate is neglected in [26], [27]. It results in a suboptimal solution. In [24], [25], the authors simplified the constrained problem into an unconstrained one, where the value of the multiplier is fixed. This leads to choosing an operating point, which may not

be the optimal one. Therefore, all these simplifications affect the SDR perceptual quality. In our case, we keep and solved the constrained problem which allows to obtain the optimal solution.

In this paper, we address two use cases. We search for a new TMO, using new gradient-based models for the HDR distortion and SDR rate, that optimizes the compression performances of the HDR backward compatible scheme. In a second time, we propose a new problem that optimizes the HDR distortion under a SDR rate constraint and under a SDR perceptual quality constraint. In a previous work [28], we demonstrated that the optimal tone mapping does not depend on the enhancement layer, therefore in this work we consider a single layer scheme.

Finally, the two tone mapping optimizations lead us to two different conclusions: (i) with the first tone mapping optimization we demonstrate that the proposed solution achieves the optimal rate distortion performance; (ii) we demonstrate that the second tone mapping optimization always provides the best trade-off between the overall rate distortion performance and the quality of the SDR content.

The remainder of this paper is organized as follows. Section II presents the compression scheme for the two use cases and the corresponding optimization problems. The models for the rate, distortion and SDR quality constraint are developed in section III as well as the corresponding tone mapping minimizations. Finally, Section IV presents the results obtained with the two TMOs.

## II. PROBLEM STATEMENT AND TMO PARAMETERS

### A. Problem statement

In this section, two different optimization problems are introduced for backward compatible HDR compression using a single layer with metadata. The first problem is the minimization of the distortion on the reconstructed HDR signal under a rate constraint. The second minimization involves the same aspects with an additional quality constraint for the SDR signal. In both cases, the compression scheme is given in Fig. 2.

To perform the R-D optimization, we model the HDR image as a random process denoted  $X$ .  $X$  generates sequences of random variable, where each random variable models a pixel value.  $X$  is tone mapped to generate an SDR image  $Y$  encoded at a given rate  $R_{SDR}$  such that the decoded SDR image  $\hat{Y}$  is compatible with legacy SDR displays.  $E_Y$  is the error between  $Y$  and  $\hat{Y}$ . The MSE between  $Y$  and  $\hat{Y}$  is the distortion introduced on  $Y$  by the encoder at a given rate  $R_{SDR}$ . To reconstruct the HDR image  $\hat{X}$ , from the decoded SDR image  $\hat{Y}$ , an inverse tone mapping is performed. Since no enhancement layer is considered, the HDR distortion  $D_{HDR}$ , generated by the encoding process, is now computed between  $X$  and  $\hat{X}$ . This image  $\hat{X}$  is compatible with new HDR displays.

The HDR image  $X$  is considered to be in a uniformly perceptual integer representation. So in practice,

$$X = \lfloor TF(X_{LL}, n_b) \rfloor \quad (1)$$

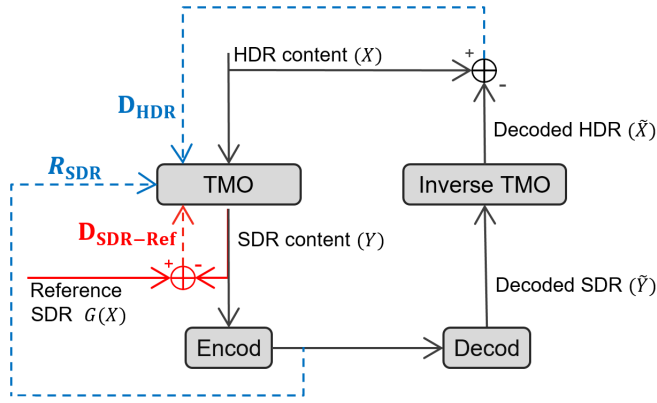


Fig. 2. Backward-compatible HDR compression scheme.

with  $X_{LL}$  the linear-light HDR signal,  $n_b$  the number of bits to represent  $X$  and  $TF$  the chosen transfer function. In this paper,  $X_{LL}$  is quantized on 12 bits with the SMPTE-2084 [8], the MPEG recommended function transfer [7]. As  $X$  is considered as a uniformly perceptual representation, the Mean Square Error (MSE) is a relevant metric to estimate HDR signals quality [21], [27]. Therefore:

$$D_{HDR} = \mathbb{E}[(\tilde{X} - X)^2] \quad (2)$$

### B. Rate-Distortion Optimization of the HDR

The goal is to find the tone mapping that minimizes  $D_{HDR}$  for a given rate budget  $R_0$ . Therefore the minimization problem is

$$\begin{aligned} \min_{TMO} \quad & D_{HDR} \\ \text{s.t.} \quad & R_{SDR} = R_0 \end{aligned} \quad (3)$$

We use an equality constraint instead of an inequality constraint since the R-D function is decreasing [29, Ex10.5] with the rate  $R$ , therefore the minimum distortion is obtained when the rate is maximal i.e. when  $R_{SDR} = R_0$ .

Problem (3) aims at minimizing the HDR distortion only. Therefore, there is no guarantee that the SDR content will have a good perceptual quality.

### C. Rate-Distortion Optimization of the HDR under a SDR quality constraint

To improve the SDR perceptual quality, in the second optimization problem, the TMO is designed to minimize the HDR distortion under a rate constraint but also under a quality constraint for the SDR signal. This constraint  $D_{SDRREF}$  is modeled as in [25] and [26]. It is a distortion between  $Y$  and an SDR image that preserves the perception and artistic intent of the HDR image. This SDR reference image is noted  $G(X)$  since it is related to the HDR image by a tone mapping curve  $G$  that has been specially designed to preserve the perception of the HDR content. Therefore,

$$D_{SDRREF} = \mathbb{E}[(G(X) - Y)^2] \quad (4)$$

With this new constraint, the problem (3) becomes

$$\begin{aligned} \min_{TMO} \quad & D_{HDR} \\ \text{s.t.} \quad & \begin{cases} R_{SDR} \leq R_0 \\ D_{SDRREF} \leq D_0 \end{cases} \end{aligned} \quad (5)$$

The contributions of the paper are twofold. First, we propose novel rate and distortion models based on the gradient of the HDR image. We show in the following that these models lead to a good approximation of the rate and of the distortion measured with a predictive codec such as HEVC [30]. Second, we derive the optimal solution of the global optimization problems (3) and (5).

### D. Piecewise Linear Tone Mapping

To solve the optimization problems (3) and (5), the TMO is approximated by a piecewise affine function (6), as in [21], [25], [26]. Let  $x$  denote a realisation of  $X$ . Then  $\forall x \in [x_k, x_{k+1}[$

$$F(x) = y = (x - x_k)s_k + y_k \quad (6)$$

where  $k \in [0, n-1[$  and where  $n$  is the number of chosen linear pieces. To ease the optimization, we aim at reducing the number of unknowns. For this reason, the interval between  $x_k$  and  $x_{k+1}$  is always the same and is noted  $\delta$ , as originally proposed in [21]. The values  $x_0, x_n$  are chosen to adaptively cover the entire range from the minimum to the maximum values of the dynamic range of the input image. Knowing  $x_0$  and  $x_n$ ,  $\delta$  is therefore fully determined with  $n$  the number of linear pieces. The minimum and maximum values of  $y$  are chosen to span the whole range of the SDR output. Therefore,  $y \in [0, 2^{n_b}-1]$  (with  $n_b$  the number of SDR bits) and finally,  $\forall x \in [(x_0 + \delta.k), (x_0 + \delta.(k+1))]$  the tone mapping (6) is

$$F(x) = \left( x - (x_0 + \delta.k) \right) . s_k + \delta \sum_{j=0}^{k-1} s_j \quad (7)$$

In this new representation (7), the only unknowns are the slopes  $\{s_k\}_{k=1 \dots n}$ . Moreover, we seek for a global invertible TMO. The parameterization in (7) leads to a continuous function. To add the invertibility constraint, it is further assume that the slopes  $s_k$  are strictly positive. Therefore, the optimization will be performed under the constraint that  $s_k > 0$ . Finally, the tone mapping curve can be assumed to be rather smooth. Therefore, the piecewise linear approximation is very accurate, even with a small number of slopes.

In the following section, only the luminance compression is considered but it could be extended to color by computing in a similar manner a TMO for the U and V components or separately for each RGB color channel.

## III. GRADIENT-BASED MODELS AND PROPOSED SOLUTION

### A. Gradient-Based SDR Bit-Rate Model

Different studies for Rate Control compare the H.264/ AVC Intra-Frame rate to different estimators based on either the gradient, the entropy, the variance and other measures of the



HDR image [31], [32]. In these studies, the gradient is shown to be the most reliable estimator. Similarly, in [33] it is shown that the HEVC Intra-Frame rate and the spatial gradient of the SDR image are highly correlated. For these reasons, we choose to model the SDR rate as a function of the SDR gradient  $\nabla_Y$  as

$$R_{SDR} = f(\mathbb{E}[\nabla_Y]) \quad (8)$$

In the sequel, we consider the following estimation for the spatial gradient of the image which are then shown to accurately predict the actual rate.

$$\nabla_Y = (|Y - Y'_H| + |Y - Y'_V|) \quad (9)$$

$$\nabla_Y = \min(|Y - Y'_H|, |Y - Y'_V|) \quad (10)$$

where  $Y'_H$  and  $Y'_V$  are the shifted versions of the image  $Y$  in the horizontal and vertical directions respectively. In (9), the sum of the horizontal and vertical gradient absolute values is computed. In (10) instead, the minimum between the vertical and horizontal gradients is computed.

To solve the optimization problems (3) and (5), we now need to express the SDR gradient (8) as a function of the parameters of the TMO. Since the TMO is related to the SDR luminance values  $Y$ , we first introduce  $Y$  in the gradient expression.

$$\begin{aligned} \mathbb{E}[\nabla_Y] &= \sum_{y=0}^{y_{max}} \mathbb{P}(Y = y) \cdot \mathbb{E}[\nabla_Y | Y = y] \\ \mathbb{E}[\nabla_Y] &= \sum_{y=0}^{y_{max}} \mathbb{P}(Y = y) \cdot \sum_{\nabla_y=0}^{\nabla_{y_{max}}} \nabla_y \cdot \mathbb{P}(\nabla_Y = \nabla_y | Y = y) \\ \mathbb{E}[\nabla_Y] &= \sum_{y=0}^{y_{max}} \sum_{\nabla_y=0}^{\nabla_{y_{max}}} \nabla_y(y) \cdot \mathbb{P}(\nabla_Y = \nabla_y, Y = y) \end{aligned} \quad (11)$$

where  $\nabla_Y$  stands for the random variable and  $\nabla_y$  its realization.  $\nabla_y(y)$  stands for the gradient value  $\nabla_y$  at a given  $y$  SDR luminance value. Let us assume that a pixel value  $X$  and its neighbor  $X'$  have similar values. Therefore, the same slope  $s_k$  is used to tone map these two values and

$$\nabla_Y = Y - Y' = X \cdot s_k - X' \cdot s_k = \nabla_X \cdot s_k \quad (12)$$

Therefore,  $\forall y \in [y_k, y_{k+1}]$ ,  $\forall x \in [x_k, x_{k+1}]$

$$\begin{aligned} \sum_{\nabla_y=0}^{\nabla_{y_{max}}} \nabla_y(y) \cdot \mathbb{P}(\nabla_Y = \nabla_y, Y = y) \\ = \sum_{\nabla_x=0}^{\nabla_{x_{max}}} \nabla_x(x) \cdot s_k \cdot \mathbb{P}(\nabla_X = \nabla_x, X = x) \end{aligned} \quad (13)$$

The approximation (13) holds when two neighboring pixels belong to the same bin (i.e same slope), this is in general true except for highly textured image region. In addition, it is interesting to note that the number of slopes used for the approximation of the TMO has little influence on the accuracy of the gradient computation (13). Indeed, reducing the number of slopes, will increase the chances that neighboring pixels are transformed with the same slope. Thus, it reduces the number of occurrence of approximation errors, but the value of these errors is potentially larger. On the other hand, increasing the number of slopes reduces the chances that neighboring pixels are transformed by the same slope but the neighboring slopes

values are increasingly closer, hence the approximation errors have a small amplitude.

From (11) and (13), the SDR gradient is now approximated by

$$\mathbb{E}[\nabla_Y] = \sum_{k=0}^{n-1} s_k \underbrace{\sum_{x=x_k}^{x_{k+1}} \sum_{\nabla_x=0}^{\nabla_{x_{max}}} \nabla_x(x) \cdot \mathbb{P}(\nabla_X = \nabla_x, X = x)}_{g_k} \quad (14)$$

where  $g_k$  only depends on the pdf of the HDR image and its gradient. This model is written for a generic formulation of the gradient and it is valid for different expressions of the gradient, as (9) and (10). More generally, the model is valid as long as the gradient is a sum of differences (as in (9)) or a minimum of differences (as in (10)).

To validate the model (14), we compare its value to the SDR gradient directly computed on the SDR image. This test has been carried out with 26 different HDR images and 3 different TMOs  $F(s_0..s_k)$ , where the slopes in  $F(s_0..s_k)$  are randomly drawn to yield a strictly increasing TMO that maps the whole range of the input HDR image. Each point is computed with a pair (image, TMO). The HDR images used for this test are very diverse [34]. They have different resolutions, up to 4K, different dynamic ranges, and issued from various sets (natural/animations, high/low spatial complexity). Fig-3(a), (b)

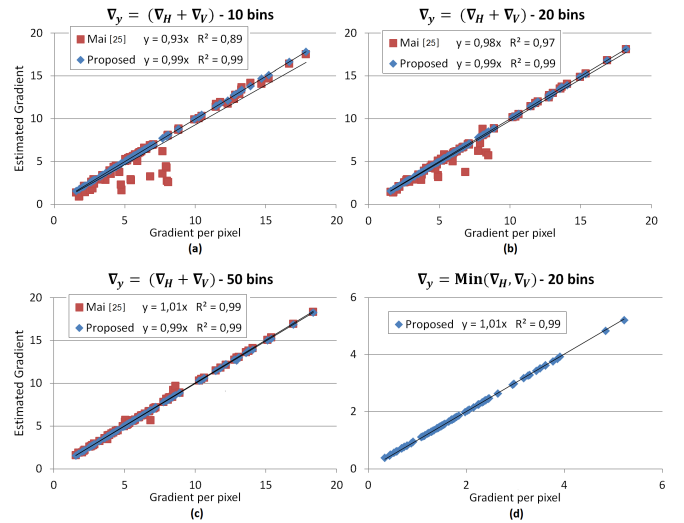


Fig. 3. Relation between SDR gradient measured on the SDR image and the values given by the proposed SDR gradient estimator (14) and the estimator used in [25] (a)-sum of the vertical and horizontal gradients (9) for 10 slopes (b)-for 20 slopes (c)-for 50 slopes (d)-min of the vertical and horizontal gradients (10) for 20 slopes

and (c) show the estimated gradient with two estimators (the proposed gradient estimator given in (9) and (14) and the one proposed in [25]) against the actual gradient values per pixel. The proposed model is more reliable whatever the number of slopes. As mentioned by the authors of [25], their estimator becomes worse when the number of slopes decreases. This aspect can be seen in Fig-3. Indeed, the estimator in [25] used a totally different model and is based on the assumption that, inside a given bin, the values of all the pixels are the same as the one of the centers of the bins. This approximation is only valid if the bin size is sufficiently small i.e. if the number

of slopes/bins is high. Therefore, their model is only accurate when using more than 50 slopes for the TMO. Fig-3(d) further shows the gradient estimated as the minimum of the horizontal and vertical gradient (10) as a function of the true gradient.

By comparing Fig-3(a-c) and Fig-3(d), we conclude that both models (9) and (10) provide a very accurate estimation of the SDR bitrate. Here, we decide to use model (10) as it mimics the selection of an intra prediction mode in a predictive coding scheme (as in HEVC-Intra). Indeed, in Intra coding, the selection of a prediction mode consists in computing the difference between the current block and a shifted version of the causal neighborhood. This looks very similar to the gradient (with only two directions: horizontal and vertical). Then, the selected mode is the one with minimum average difference value.

The scatter plots in Fig-4 compare the Intra-Frame HEVC rate of the SDR content and its estimate. The SDR rate estimate is based on the gradient (14) in Fig-4(a) and on the entropy of the SDR signal [23] in Fig-4(b). We use the same test set used in Fig-3 (26 images x 3 TMO) with 20 slopes for the TMOs. Similar results were obtained with different numbers of slopes. Each point represents a pair (image, TMO).

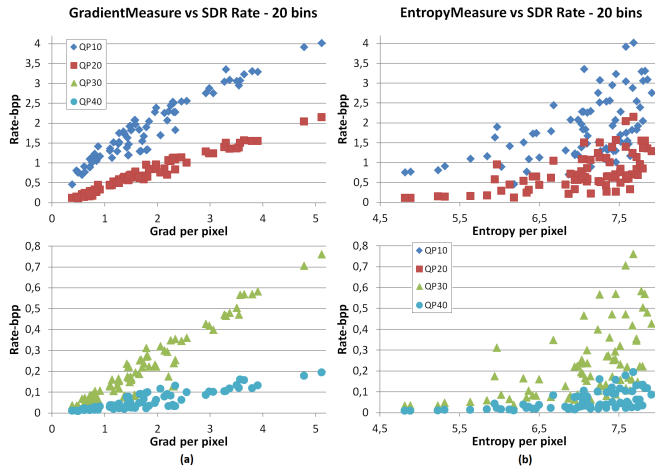


Fig. 4. Experimental validation of the proposed rate model (a)-Correlation between Gradient-based estimator and rate for QP = [10, 20, 30, 40] (b)-Correlation between Entropy-based estimator and rate for QP = [10, 20, 30, 40]

Fig-4(a) shows less dispersion than the scatter plots in Fig-4(b). Therefore, the HDR image gradient in Fig-4(a) provides a more accurate estimate of the actual rate, for all QP (Quantization parameter in HEVC), than the entropy in Fig-4(b). Moreover, an affine function seems sufficient to model the relationship between the HDR image gradient and the rate of any images. The SDR rate function (8) can be rewritten as:

$$R_{SDR} = a \cdot \mathbb{E}[\nabla_Y] + b \quad (15)$$

$$= a \cdot \sum_{k=0}^{n-1} (s_k \cdot g_k) + b \quad (16)$$

Fig-5 represents the most appropriate  $a$  and  $b$  values for (16) to approximate the cloud points in Fig-4(a). Using these values allows us to find the SDR rate of any image, knowing the

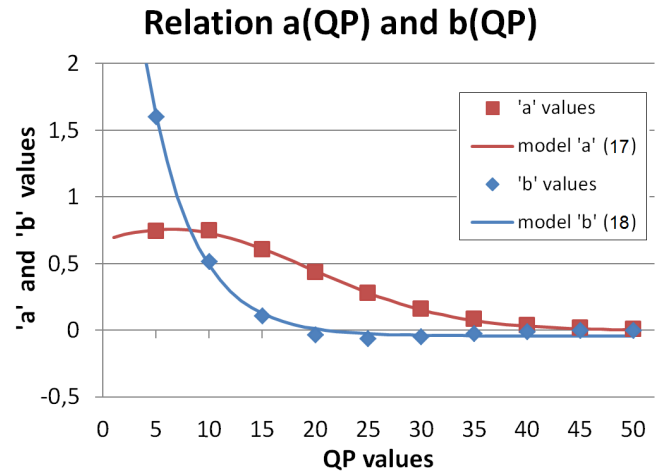


Fig. 5. Values of coefficients 'a' and 'b' versus QP and related models

HDR image gradient  $g_k$  and the applied TMO  $s_k$ . Moreover, a relationship between the QPs and the coefficients 'a' and 'b' have been obtained through nonlinear least squares fitting (Trust-Region algorithm [35]).

$$\hat{a} = 0,7567 \cdot e^{-\left(\frac{QP-6,337}{18,72}\right)^2} \quad (17)$$

$$\hat{b} = 5,161 \cdot e^{-0,228 \cdot QP} - 0,044 \quad (18)$$

The model (16) and these coefficients allow to estimate the rate of any HDR image using only its gradient and the chosen TMO.

### B. Classical HDR Distortion Model

Considering the MSE in the chosen transfer domain, the HDR distortion in [21], [23], [25], [26], [28] is

$$D_{HDR} = \mathbb{E}[(\tilde{X} - X)^2] \quad (19)$$

In [21], [25], [26], using the parameterization (6), the HDR distortion becomes

$$D_{HDR} = \sum_{k=0}^{n-1} \sum_{x=x_k}^{x_{k+1}} \sum_{\tilde{y}=\tilde{y}_k}^{\tilde{y}_{k+1}} (\tilde{x}(\tilde{y}, s_k) - x)^2 \cdot \mathbb{P}(E_Y = (\tilde{y} - y), X = x) \quad (20)$$

[21], [23], [25], [26], [28] assume that the quantization error on the SDR signal  $E_Y = \tilde{Y} - Y$  is independent of the SDR image  $Y$  (called *independence assumption*). Under this assumption, the HDR distortion becomes

$$D_{HDR} = \sigma_{E_Y}^2 \cdot \mathbb{E}[F'(X)^{-2}] \quad (21)$$

where ' stands for the first derivative and where  $\sigma_{E_Y}^2$  is the variance of the SDR signal error  $E_Y$ . Using the parameterization (6) of [21], [25], [26], the HDR distortion (21) becomes

$$D_{HDR} = \sigma_{E_Y}^2 \cdot \sum_{\forall k} s_k^{-2} \cdot \mathbb{P}(x_k \leq X < x_{k+1}) \quad (22)$$

(22) leads to an accurate estimation as shown in Fig-6. Indeed, the scatter plots shows a high correlation between the true HDR distortion and the estimated one (22). This experimental

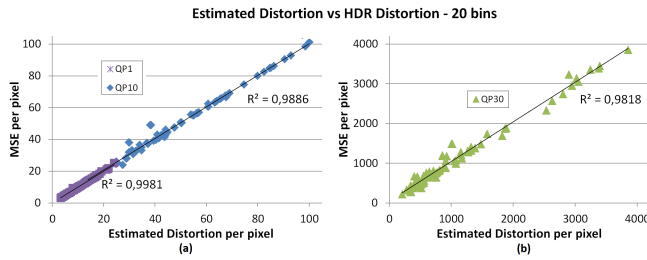


Fig. 6. Experimental validation of the estimated distortion (22) (a)-At very-high rate (QP1) and high rate (QP10) (b)-At medium-low rate (QP30)

validation use the same test set as in Fig-3 (26 images, 3 TMOs).

However, (22) is intractable in practice since to evaluate the variance  $\sigma_{E_Y}^2$ , one needs to compress the data and get  $\tilde{Y}$ : it's a vicious circle. In order to overcome this difficulty, [23], [25], [26] assume that  $\sigma_{E_Y}^2$  depends on the QP factor only but does not depend on the statistics of the SDR image (called *invariance assumption*). In [21], they assume that  $\sigma_{E_Y}^2$  is multiplicative constant and neglect it. Thus, in both cases,  $D_{HDR}$  is rewritten as:

$$D_{HDR} \propto \sum_{\forall k} s_k^{-2} \cdot \mathbb{P}(x_k \leq X < x_{k+1}) \quad (23)$$

To test the validity of the independence and the statistics invariance assumptions, we compute the estimated distortion expressed in (23) and compare it to the corresponding HDR distortion at different QPs, as in Fig-6. Fig-7 shows that the

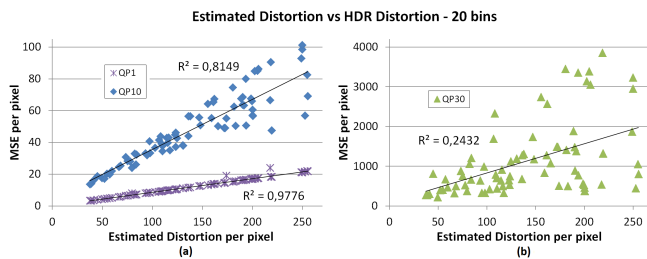


Fig. 7. Experimental validation of the estimated distortion (23) (a)-At very-high rate (QP1) and high rate (QP10) (b)-At medium-low rate (QP30)

distortion  $D_{HDR}$  scales linearly with the estimated  $D_{HDR}$  (23) at low QP only. At QP=10, the correlation coefficient is much lower and it decreases as QP increases. At QP=30, there is little correlation between the measured and estimated values. Fig-7 shows the distortion when the TMO use 20 slopes but, as in Fig-3, the results were the same with 50 slopes. These observations are mainly due to the assumed independence between the SDR image  $Y$  and its compression error  $E_Y$ , which is not valid as shown in Fig-8. Indeed, the error  $E_Y$  still contains details of the original image  $Y$ , especially for high QP.

### C. Proposed Gradient-Based HDR Distortion Model

As in [21], [23], [25], [26], [28] the proposed HDR distortion model is a MSE in the chosen transfer domain. Therefore,

using the parameterization (6), we use the expression (20) for the HDR distortion. Here we neither use the independence nor the invariance assumption, instead, the proposed formulation takes into account the joint distribution of the HDR image  $X$  and the compression error  $E_Y$ .

Considering that the luminance value of the original and reconstructed SDR image ( $y$  and  $\tilde{y}$  respectively) belong to the same bin  $[x_k, x_{k+1}[$  and therefore the same slope  $s_k$ , the HDR distortion (20) becomes:

$$D_{HDR} = \sum_{k=0}^{n-1} \sum_{x=x_k}^{x_{k+1}} \sum_{\tilde{y}=\tilde{y}_k}^{\tilde{y}_{k+1}} \frac{(\tilde{y} - y(x, s_k))^2}{s_k^2} \cdot \mathbb{P}(E_Y = (\tilde{y} - y), X = x) \quad (24)$$

$$D_{HDR} = \sum_{k=0}^{n-1} \frac{1}{s_k^2} \cdot \sum_{x=x_k}^{x_{k+1}} \sum_{\tilde{y}=\tilde{y}_k}^{\tilde{y}_{k+1}} E_y^2(\tilde{y}, x, s_k) \cdot \mathbb{P}(E_Y = (\tilde{y} - y), X = x) \quad (25)$$

where  $E_y(\cdot)$  stands for the function that computes the SDR error from the HDR content  $x$ , the reconstructed SDR  $\tilde{y}$  and the TMO slope  $s_k$ .

Fig-8 shows that the compression error  $E_Y$  depends on the gradient of the SDR signal and that this dependence increases with the QP value. We propose to model  $E_Y^2$  as an affine function of  $(\nabla_Y)^\gamma$ , where the coefficients of the affine function and  $\gamma$  depend on the QP value only. This leads to a new model for  $D_{HDR}$  where (25) becomes

$$D_{HDR} \propto \sum_{k=0}^{n-1} \frac{1}{s_k^2} \cdot \sum_{x=x_k}^{x_{k+1}} \sum_{\nabla_y=0}^{\nabla_{ymax}} (\nabla_y(y(x, s_k)))^\gamma \cdot \mathbb{P}(\nabla_Y = \nabla_y, X = x) \quad (26)$$

Following the same assumption as in (12) and (13), we obtain

$$D_{HDR} \propto \sum_{k=0}^{n-1} \frac{g_k(\gamma)}{s_k^{2-\gamma}} \quad (27)$$

where

$$g_k(\gamma) = \sum_{x=x_k}^{x_{k+1}} \sum_{\nabla_x=0}^{\nabla_{xmax}} (\nabla_x(x))^\gamma \cdot \mathbb{P}(\nabla_X = \nabla_x, X = x) \quad (28)$$

Note that the  $g_k(\gamma)$  values depend on the HDR image and  $\gamma$  only.

To complete the HDR distortion model, we need to estimate the relationship between  $\gamma$ , the affine coefficients and QP. We first consider the estimation of  $\gamma$ . With sufficient image tests (same test set as in Fig-3), we compare the estimated distortion (27), computed with different  $\gamma$  values, to the corresponding HDR distortion at a given QP. For each QP value, we retain the  $\gamma$  value which maximizes the correlation coefficient between the distortion and its estimation. The best  $\gamma$  values are plotted in Fig-9(a). In the worst case, the correlation coefficient is 0.945 at QP16.

We now establish a relationship between the affine coefficients  $c$  and  $d$  of the HDR distortion model (29) and QP.

$$D_{HDR} = c \cdot \sum_{k=0}^{n-1} \left( \frac{g_k(\gamma)}{s_k^{2-\gamma}} \right) + d \quad (29)$$



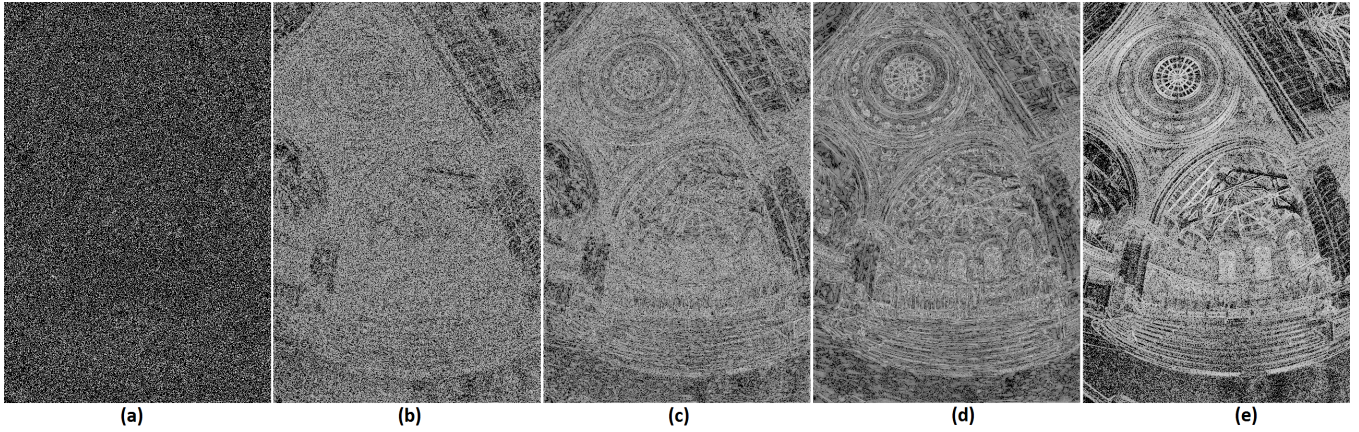


Fig. 8. Error  $E_Y$  between SDR source and reconstructed SDR compared to SDR source gradient  $\nabla_Y$  (a)-error at QP=10 (b)-error at QP=20 (c)-error at QP=30 (d)-error at QP=40 (e)-Gradient – For display and print purposes, the displayed images are modified as follows:  $(E_Y^2)^{1/8}$  and  $(\nabla_Y)^{1/8}$

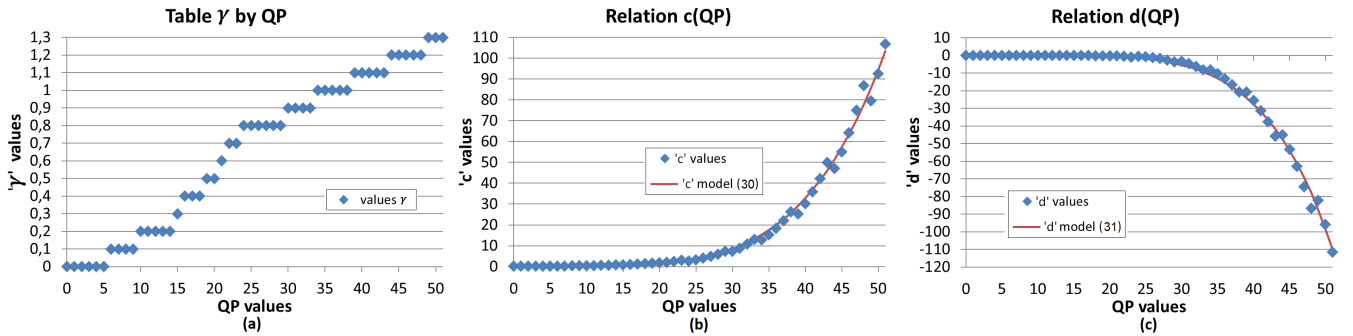


Fig. 9. (a)-  $\gamma$  values versus QP (b)- values of coefficient  $c$  versus QP and related model (c)- values of coefficient  $d$  versus QP and related model

As shown in Fig-9(b) and (c), the relationship between  $c$ ,  $d$  and the QP can be modeled using a nonlinear least squares fitting (Trust-Region algorithm [35]) as:

$$\hat{c} = (8.939 \cdot 10^{-7}) \cdot QP^{4.722} + 0.124 \quad (30)$$

$$\hat{d} = (-2,223 \cdot 10^{-8}) \cdot QP^{5.677} + 0.034 \quad (31)$$

Finally, the parameters of the functions (30-31) and a table for  $\gamma$  (one value per QP is stored leading to 51 entries) are stored at the encoder to perform the TMO. Since these values are used at the encoder only, there is no need to send them.

As we made this learning process with different numbers of slopes and that the resulting values were the same, we conclude that  $\gamma$ ,  $c$  and  $d$  relations are valid for any number of slopes. Fig-10 shows that the proposed distortion estimator (27) is better than the classical model (23) represented in Fig-7. Based on the well correlated offline models for  $c$ ,  $d$  and  $\gamma$  parameters, the proposed estimator (29) has the advantage to estimate the distortion for any input HDR image; while in the classical estimator (22)  $\sigma_{E_Y}^2$  remains unknown and dependent of the image.

Note that, although the proposed distortion model is here applied in the context of HDR encoding, it could be used to predict the encoder distortion at a given rate using only the SDR image and its gradient.

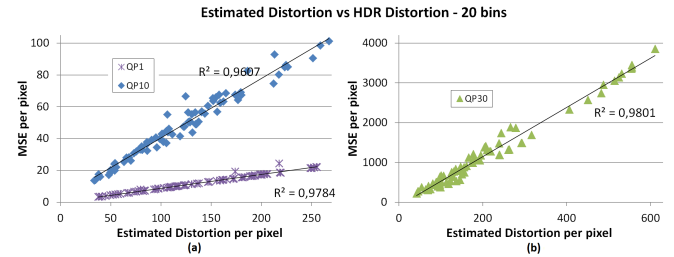


Fig. 10. Experimental validation of the distortion estimated with the proposed model (a)-At very-high rate (QP1) and high rate (QP10) (b)-At medium-low rate (QP30)

#### D. Rate-Distortion Optimization

We now aim at solving the optimization problem (3). As explained in Section-II-D, the TMO is chosen to span the whole range of the SDR output. Therefore, we need to add a new constraint to the minimization problem to obtain an image in the specified SDR range. Moreover, the model for  $R_{SDR}$  (16) and  $D_{HDR}$  (29) introduce the hidden parameter  $QP$ . Therefore, (3) requires an additional optimization over this hidden parameter  $QP$ .

$$\min_{QP} \min_{s_1 \dots s_n} c \cdot \sum_{k=0}^{n-1} \left( \frac{g_k(\gamma)}{s_k^{2-\gamma}} \right) + d \quad (32a)$$



$$\text{s.t.} \quad \begin{cases} \sum_{k=0}^{n-1} s_k \cdot g_k(1) = \frac{R_0 - b}{a} \\ \sum_{k=0}^{n-1} s_k = \frac{2^{n_b-1}}{\delta} \end{cases} \quad (32b)$$

$$(32c)$$

where  $g_k(\gamma)$  is defined in (28),  $a$  and  $b$  in (17-18),  $c$  and  $d$  in (30-31),  $n_b$  and  $\delta$  in Section-II-D.

The optimization over the QP parameter is either solved exhaustively or by learning a relationship between QP and the rate constraint  $R_0$ , see Section IV-A for details. We now solve the second minimization problem over the TMO parameters  $\{s_k\}$  and assume that the QP parameter is fixed. The Karush-Kuhn-Tucker necessary conditions for optimality [36, Chap5.5.3] on the Lagrangian expression of (32) lead to:

$$\begin{cases} \frac{-(2-\gamma)g_j(\gamma).c}{s_j^{(3-\gamma)}} + \lambda.g_j(1) + \mu = 0 \quad \forall j \in [0, n] \end{cases} \quad (33a)$$

$$\begin{cases} \sum_{k=0}^{n-1} (s_k \cdot g_k(1)) - \frac{R_0 - b}{a} = 0 \end{cases} \quad (33b)$$

$$\begin{cases} \sum_{k=0}^{n-1} (s_k) - \frac{2^{n_b-1}}{\delta} = 0 \end{cases} \quad (33c)$$

Fortunately, (33a) depends on one  $s_j$  value only. Therefore, the optimal solution of (32) is

$$s_k = \left( \frac{(2-\gamma).g_k(\gamma).c}{\mu + \lambda.g_k(1)} \right)^{\frac{1}{3-\gamma}} \quad (34)$$

where  $\mu$  and  $\lambda$  are the solutions of:

$$\begin{cases} \sum_{k=0}^{n-1} \left( \frac{(2-\gamma).g_k(\gamma).c}{\mu + \lambda.g_k(1)} \right)^{\frac{1}{3-\gamma}} \cdot g_k(1) = \frac{R_0 - b}{a} \end{cases} \quad (35a)$$

$$\begin{cases} \sum_{k=0}^{n-1} \left( \frac{(2-\gamma).g_k(\gamma).c}{\mu + \lambda.g_k(1)} \right)^{\frac{1}{3-\gamma}} = \frac{2^{n_b-1}}{\delta} \end{cases} \quad (35b)$$

where  $g_k(\gamma)$  is defined in (28),  $a$  and  $b$  in (17-18),  $c$  and  $d$  in (30-31),  $n_b$  and  $\delta$  in Section-II-D. The system (35) is solved numerically. The results of this minimization are discussed in Section-IV

Note that, one can find a suboptimal but analytical solution to (32) by omitting the rate constraint (32b), we only minimize the distortion and find the following analytical solution:

$$s_k^\circ = \frac{2^{n_b-1} \cdot g_k(\gamma)^{\frac{1}{3-\gamma}}}{\delta \cdot \sum_{k=0}^{n-1} g_k(\gamma)^{\frac{1}{3-\gamma}}} \quad (36)$$

On the other hand, if we remove the spanning constraint (32c), we obtain:

$$s_k^\dagger = \alpha \cdot \frac{g_k(\gamma)^{\frac{1}{3-\gamma}}}{g_k(1)^{\frac{1}{3-\gamma}} \cdot \sum_{k=0}^{n-1} (g_k(1)^{\frac{2-\gamma}{3-\gamma}} \cdot g_k(\gamma)^{\frac{1}{3-\gamma}})} \quad (37)$$

where  $\alpha$  is adjusted to respect the removed spanning constraint. The derivation of these two TMOs is detailed in the supplementary documents [34].

**Remark.** Experimentally, we observe that the TMO curve defined by  $\{s_k\}$  is always between the TMO curves defined by  $\{s_k^\dagger\}$  and  $\{s_k^\circ\}$  respectively. This observation also holds when the TMO curves defined by  $\{s_k^\dagger\}$  and  $\{s_k^\circ\}$  intersect.

Note that the solution proposed in [21] is a particular case of equation (36). More precisely, the TMO in [21] is the same

as our high rate regime TMO, computed for  $\gamma = 0$  (and thus  $\nabla_x^\gamma = 1$ ,  $g_k(0) = \sum_{x=x_k}^{x_{k+1}} \mathbb{P}(X = x)$ ).

### E. R-D Optimization with SDR quality constraint

We now consider the use of an additional SDR quality constraint, i.e., the optimization problem given by (5). The models for  $D_{HDR}$  and  $R_{SDR}$  have been expressed in the previous section, however we also need to model the SDR constraint quality  $D_{SDRREF}$ .

As in [25], [26], the SDR quality compared with a reference tone-mapped SDR is computed as:

$$D_{SDRREF} = \mathbb{E}[(G(X) - Y)^2] \quad (38)$$

$$D_{SDRREF} = \sum_{k=0}^{n-1} \sum_{x=x_k}^{x_{k+1}} (G(x) - F(x))^2 \cdot \mathbb{P}(X = x) \quad (39)$$

with  $G$  a global TMO optimized for the SDR perceptual quality and  $F$  the TMO defined in (6). In [25], [26], a coarse assumption is made:

$$\forall x \in [x_k, x_{k+1}[ , \quad x = \frac{x_{k+1} - x_k}{2} \quad (40)$$

which yields a correct result if the function to integrate  $(G(x) - F(x))^2$  is linear in  $x$  and if the distribution is uniform on  $[x_k, x_{k+1}[$ . This leads to a piecewise constant distortion measure. Instead, we propose to keep only the uniform distribution assumption (41).

$$\forall x \in [x_k, x_{k+1}[ , \quad \mathbb{P}(X = x) = \frac{\mathbb{P}(x_k \leq X < x_{k+1})}{\delta} \quad (41)$$

Then (39) becomes:

$$D_{SDRREF} = \sum_{k=0}^{n-1} \frac{p_k}{\delta} \cdot \sum_{x=x_k}^{x_{k+1}} (G(x) - F(x))^2 \quad (42)$$

where  $p_k = \mathbb{P}(x_k \leq X < x_{k+1})$ .

If we parameterize the function  $G$ ,  $\forall x \in [x_k, x_{k+1}[$ , with

$$G(x) = ((x - x_k).t_k + z_k) \quad (43)$$

The expression (42) becomes

$$D_{SDRREF} = \sum_{k=0}^{n-1} p_k \cdot h(t_0 \dots t_k, s_0 \dots s_k, \delta) \quad (44)$$

with

$$h(t_0 \dots t_k, s_0 \dots s_k, \delta) = \frac{(\delta.t_k + z_k - \delta.s_k - y_k)^3 - (z_k - y_k)^3}{\delta \cdot 3 \cdot (t_k - s_k)} \quad (45)$$

The derivation of (44) is detailed in the supplementary documents [34].

To test the validity of the assumption (41), we compare the MSE described in (39) and the estimated distortion expressed in (44) in Fig-11. This test has been carried out with 26 different images and 3 different TMOs  $F(s_0 \dots s_k)$  (same test set as in Fig-3), where the slopes in  $F(s_0 \dots s_k)$  are randomly drawn to yield a strictly increasing TMO that maps the whole range of the input HDR image. In this test, the reference TMO

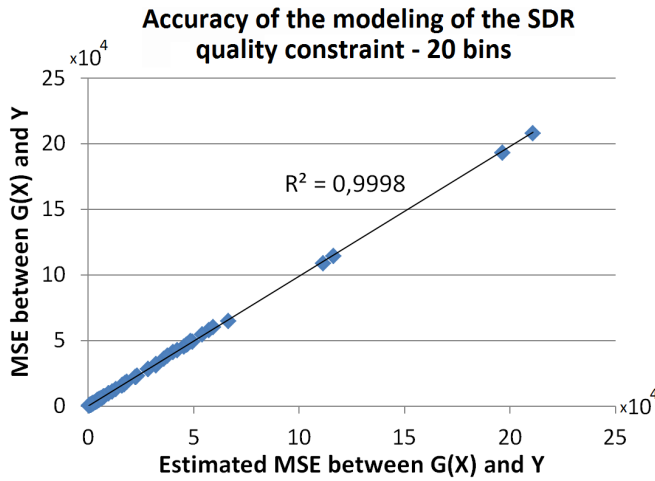


Fig. 11. Experimental validation of the estimated SDR quality constraint with (44)

$G(t_0..t_k)$  is always the same global TMO [19] approximated with  $n$  slopes. The correlation between the estimated and real values (39) is very high, demonstrating that the assumption made is very acceptable. Fig-11 shows the distortion when we use 20 slopes for both TMOs. As in Fig-3, the results were the same with 50 slopes. This test is also performed with the SDR quality constraint proposed in [25], the results are the same. One possible explanation of the close results obtained for both  $D_{SDRREF}$  expression is the use of very different tone mapping curves for the comparison. We suppose that with the use of closer tone mapping curves, expression in [25] leads to approximation errors for  $D_{SDRREF}$ .

With the developed models for  $R_{SDR}$  (16),  $D_{HDR}$  (29),  $D_{SDRREF}$  (44) and with the spanning constraint, the minimization problem (5) becomes

$$\begin{aligned} \min_{QP} \min_{s_1 \dots s_n} \quad & c \cdot \sum_{k=0}^{n-1} \left( \frac{g_k(\gamma)}{s_k^{2-\gamma}} \right) + d \\ \text{s.t.} \quad & \begin{cases} \sum_{k=0}^{n-1} s_k \cdot g_k(1) \leq \frac{R_0 - b}{a} \\ \sum_{k=0}^{n-1} s_k = \frac{2^{n_b-1}}{\delta} \\ \sum_{k=0}^{n-1} p_k \cdot h(t_0 \dots t_k, s_0 \dots s_k, \delta) \leq D_0 \end{cases} \end{aligned} \quad (46)$$

This problem can be solved numerically with the interior-point optimization method. The results of this minimization are discussed in Section-IV

#### IV. EXPERIMENTAL RESULTS

##### A. R-D Optimization

In this section, we show the RD performances achieved with the backward compatible HDR compression scheme of Figure 2 with the TMO curve obtained in Section III-D. As in the previous experiments, the TMOs are computed with 20 slopes. Tests were made with 20 and 50 slopes and the RD performances were the same.

The proposed TMO design considers one TMO per frame hence assumes the TMO parameters to be sent for each frame. This cost has been counted in the rate cost. However, the rate

cost for transmitting the TMO parameters is negligible. Using 16 bits for each of our 20 slopes leads to an extra rate of 0.0002 bits/pixel for a 1920x1080 image ( $\frac{16 \cdot 20}{1920 \cdot 1080}$ ). This cost can be further reduced by considering entropy coding of the slopes.

The original HDR image is first perceptually transformed using the SMPTE-2084 PQTF to follow the contrast sensitivity function of the human eye, and uniformly quantized to 12 bits. This provides the HDR content denoted  $X$ .  $X$  is then tone mapped using the TMO that minimizes (32) and the resulting SDR content is encoded with the HEVC reference software (HM 16.2). First, the optimal TMO is derived. Rather than solving Problem (32), where an exhaustive search over the QP parameter for a given rate constraint  $R_0$ , we solve an equivalent problem, where an exhaustive search over the rate constraint  $R_0$  is performed for a given QP parameter.

---

##### Algorithm 1 RD simulation: convex hull

---

**Require:** HDR image  $X$

**for all** QP **do**

    Compute  $\nabla_X$ , then  $p(X, \nabla_X)$

    Compute  $a, b, c, d, \gamma$  and  $g_k(\gamma)$  from (17), (18), (30), (31), Fig-9(a) and (28) respectively.

    Compute the two TMO bounds  $\{s_k^\dagger\}$  (36) and  $\{s_k^\circ\}$  (37)

    For each bound, compute the necessary rate:  $R_{SDR}^\dagger$  and  $R_{SDR}^\circ$  from (16)

    Quantize the interval  $[R_{SDR}^\dagger, R_{SDR}^\circ]$  to get 20 values.

**for all**  $R_0 \in [R_{SDR}^\dagger, R_{SDR}^\circ]$  **do**

        Solve numerically (35) to get  $\lambda, \mu$

        Compute  $\{s_k\}$  with (34)

        Apply the TMO (defined by  $s_k$ ) to  $X \rightarrow$  SDR image

        Encode the SDR image  $\rightarrow R_{SDR}$

        Inv. Tone Mapped the decoded SDR image

        Compute  $D_{HDR}$

**end for**

**end for**

    Choose the best  $R_{SDR} - D_{HDR}$  point

---

Note that in Algorithm 1, the search over the optimal  $R_0$  is performed in a reduced interval  $[R_{SDR}^\dagger, R_{SDR}^\circ]$ . This is possible because the TMO curves (36) and (37) are obtained without the spanning and rate constraints respectively and therefore lead to lower and upper bounds on the required encoding rate.

Fig-12 shows the RD performance obtained with Algorithm 1. For each QP value, a RD curve parameterized by  $R_0$  shows the RD performance obtained with the optimized TMO. Note that the left-most and right-most points of each curve at QP fixed correspond to the rates  $R_{SDR}^\dagger$  and  $R_{SDR}^\circ$ . Then the convex hull of all these curves is drawn. The optimal rate constraints  $R_0$  correspond to the case, where each individual curve (QP fixed) is tangent to the convex hull.

The exhaustive search in Algorithm-1 is rather complex as it requires encoding the SDR image for each  $R_0$  value. However, the loop over  $R_0$  can be avoided by learning a correspondence between this rate constraint  $R_0$  and the QP parameter. To learn such a relation, an exhaustive search is performed for multiple images such that we obtain the optimal

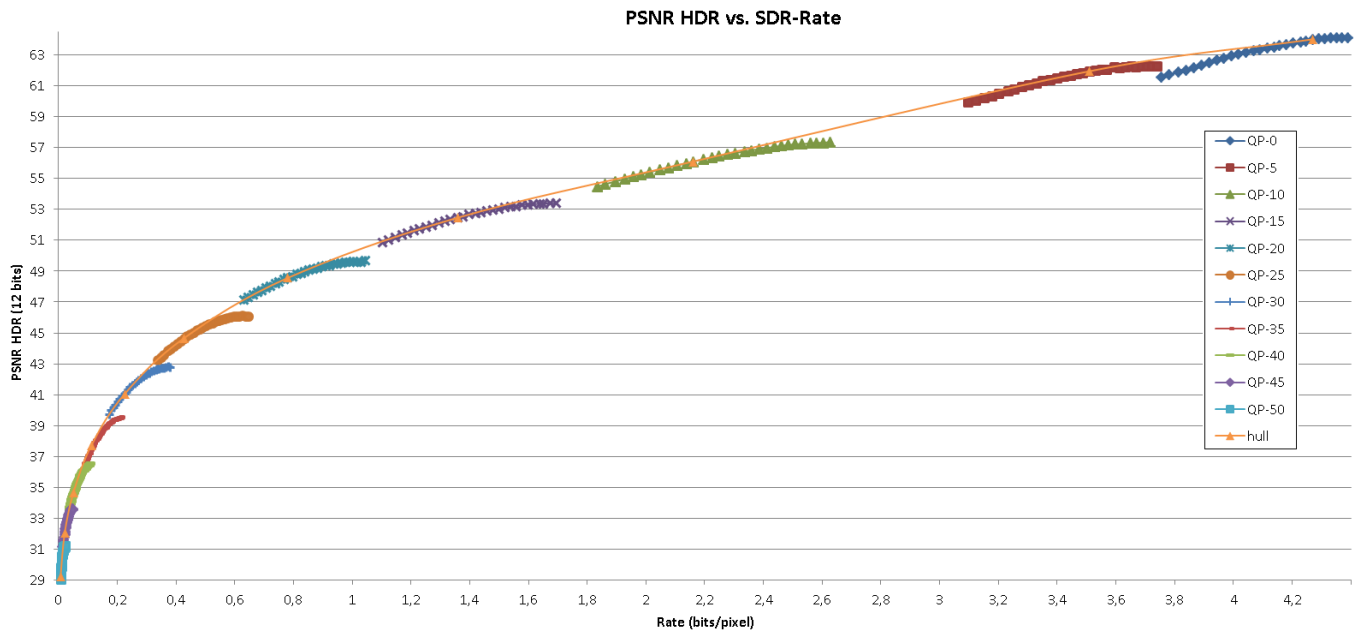


Fig. 12. RD optimization without SDR quality constraint Algorithm-1. Given a QP value, different  $R_0$  values are tested. The best RD performance is given by the convex envelop of all the points.

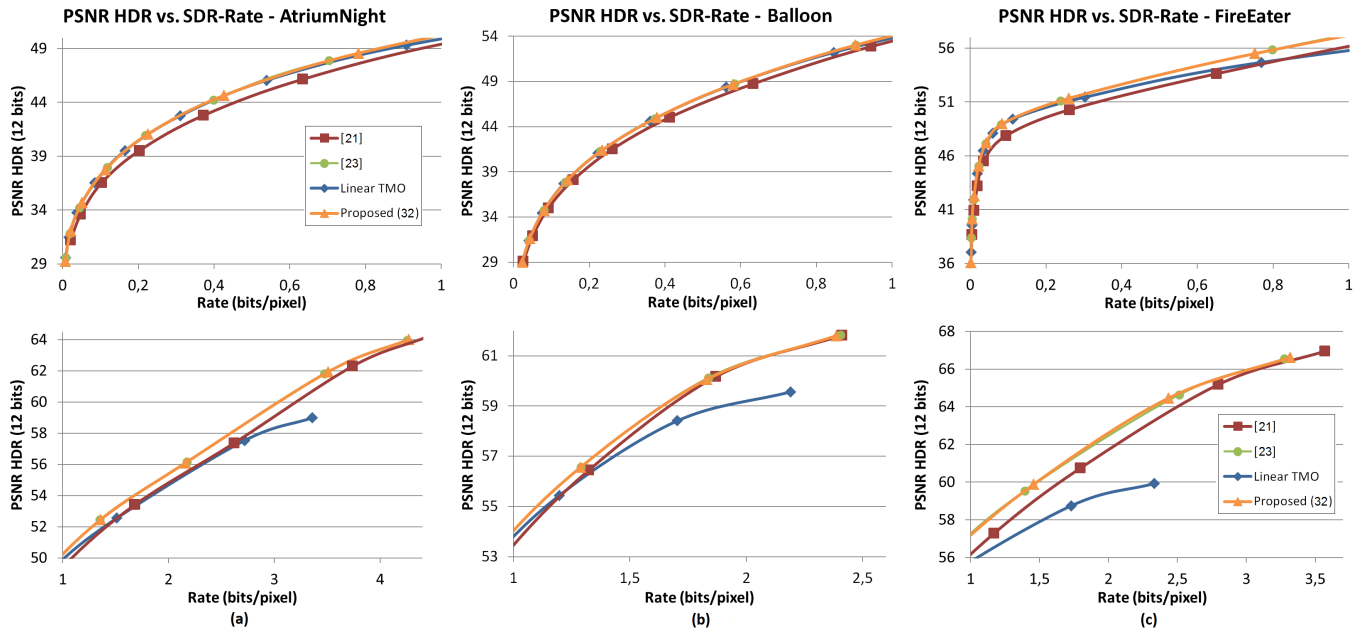


Fig. 13. Distortion between the original HDR image  $X$  and the reconstructed image  $\tilde{X}$  vs SDR rate - (a) Results for image AtriumNight at low and high rates (b) Results for image Balloon at low and high rates (c) Results for image FireEater at low and high rates

$R_0$  values for each QP. Then, a law is fitted to these optimal  $R_0$  points. Given a QP value, at run time, this law allows us to have directly the optimal  $R_0$  value. Deducing the optimal TMO then becomes feasible in real-time. Fig-13 shows the RD performances obtained with this straightforward method, in comparison with several existing TMOs: a simple linear TMO, a TMO that minimizes the distortion only [21] and a TMO optimized according to a RD criterion, but with a RD model valid in the high rate regime only [23]. The experiment is done with images from the MPEG test set and the one

proposed in [37], [38]. Fig-14 shows different tested images. More images are tested in the supplementary materials [34]. The upper charts show the PSNR of the reconstructed HDR image versus the SDR image bitrate for high QPs (0 to 1 bit/pixel) while the lower ones show the same results for lower QPs.

First, we notice that, as expected, the proposed algorithm, performing a RD optimization, allows to outperform the linear and [21] TMOs at any rate. More precisely, the Bjontegaard rate gains between the proposed and the linear TMOs are 3.0%

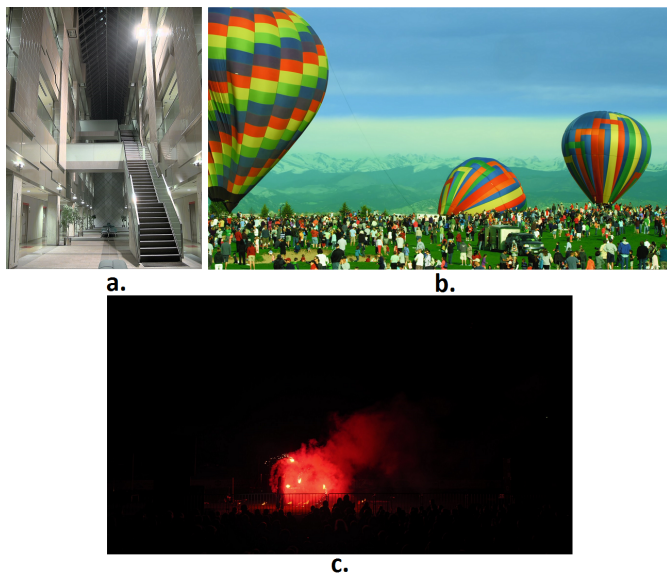


Fig. 14. Test HDR images. For display purposes, the HDR images are tone-mapped with [19] : a) AtriumNight b) Balloon c) FireEater

|                   | AtriumNight | Balloon  | FireEater |
|-------------------|-------------|----------|-----------|
| RD Gain to [21]   | 15.5%       | 7.5%     | 26.9%     |
| RD Gain to Linear | 3.0%        | 1.4%     | 11.4%     |
| Dynamic Range     | $4.10^9$    | $9.10^7$ | $7.10^6$  |
| Spatial Activity  | 76          | 82       | 38        |

Fig. 15. RD Gains, Dynamic Range and Spatial Activity of the tested images as defined in [39]

for AtriumNight, 11.4% for FireEater and 1.4% for Balloon. The Bjontegaard rate gains between the proposed TMO and [21] are 15.5% for AtriumNight, 26.9% for FireEater and 7.5% for Balloon. At high rates, the TMO always seems to converge towards the same R-D performances as [21] and towards the same R-D performances as the linear TMO at low rates. For each image, the R-D performances are different but the behavior seems to be the same regardless of the dynamic range or spatial activity (defined in [39]), as one can see in Fig-15.

Regarding [23], the performances are very close to the proposed method for all the tested images. This result is surprising since we demonstrated the better accuracy of the proposed model for the rate and distortion in Section-III-A and Section-III-C. However, [23] learns the Lagrangian multiplier in the RD optimization. This learning compensates for the independence and invariance assumptions (see Section-III-C) made in their distortion model and leads to a corrected RD model, which is as accurate as our. This shows indeed how important the accuracy of the RD model is.

### B. R-D Optimization with SDR quality constraint

We now analyze the RD performance obtained with the TMO solution of (46) with a new constraint on the SDR perceptual quality. The problem is solved with the interior point algorithm (optimization toolbox from Matlab). The resulting TMO depends on two parameters: the  $D_{SDRREF}$

constraint  $D_0$  and the rate constraint  $R_0$ .  $D_0$  is adjusted to 34.2dB (MSE=25) to insure sufficient similarity with the SDR reference image.

The rate constraint  $R_0$  needs to be higher than in the previous Section IV-A since adding a new constraint will necessarily increase the required rate. Second, the rate constraint will depend on the reference Tone-Mapping chosen. For instance, Figure-16 plots the bounding TMOs described in section-IV-A and the chosen reference Tone-Mapping [19], the PTR (Photographic tone reproducer). For AtriumNight and Balloon images, the PTR is close to the two bounding TMOs, whereas the PTR is quite far from these bounds for FireEater and this will significantly increase the rate (since we observed that the optimal TMO with best RD trade off lies in between the two bounds). Therefore, to find the optimal TMO for a given QP parameter, a first rate constraint  $R_{max}$  is computed as the maximum between the rate needed with the PTR TMO and the rate needed for  $\{s_k^\dagger\}_{k=0..n}$  (36), i.e. the upper bound on the rate constraint computed in the previous Section IV-A.  $R_{max}$  ensure to find a solution to (46). Using the previously learnt relationship between  $R_0$  and the QP parameter for the RD optimization without SDR constraint (section IV-A), we can compute the optimal rate constraint  $R_0$ . If the optimal  $R_0$  does not satisfy the SDR quality constraint when solving (46),  $R_0$  is increased until we reach  $R_{max}$ . The loop stops when we find the smallest  $R_0$  value that satisfies the SDR constraint. The complexity of this solution is fast: under 1 sec with the Matlab optimization toolbox.

The R-D performance of this TMO is compared with two state of the art TMO designed for R-D performances under a quality constraint for the SDR image [25], [26] and with the unconstrained one defined in the previous section. With this test, the R-D losses induced by the SDR quality constraint can be measured. Experimental test conditions are identical to the previous section. The results are shown on Fig-17.

The left and center charts plot the PSNR of the reconstructed HDR image versus the SDR image bitrate at low and high rate respectively. The right charts plot the PSNR of the SDR constraint quality  $D_{SDRREF}$  versus the SDR bitrate. The proposed tone mapping saves 11.78% 39.17% and 2% of bitrate for the image AtriumNight, FireEater and Balloon compared to [26] and 5.44%, 1.95% and 2.66% compared to [25] for the same or even better SDR quality constraint. This can be explained by the fact that our proposed RD model is more accurate than those in [25] and [26]. Moreover, as expected from Fig-16, the SDR quality constraint has a much bigger impact on the image FireEater.

## V. CONCLUSION

This paper presented two new TMOs for HDR backward compatible compression. Using the HDR image gradient, we provided new statistical models for estimating  $D_{HDR}$ ,  $R_{SDR}$  and  $D_{SDRREF}$  and showed their accuracy to the real data. The first TMO minimizes the distortion of the HDR image under a rate constraint on the SDR layer. The second TMO remains the same minimization with an additional constraint to preserve the SDR perceptual quality.



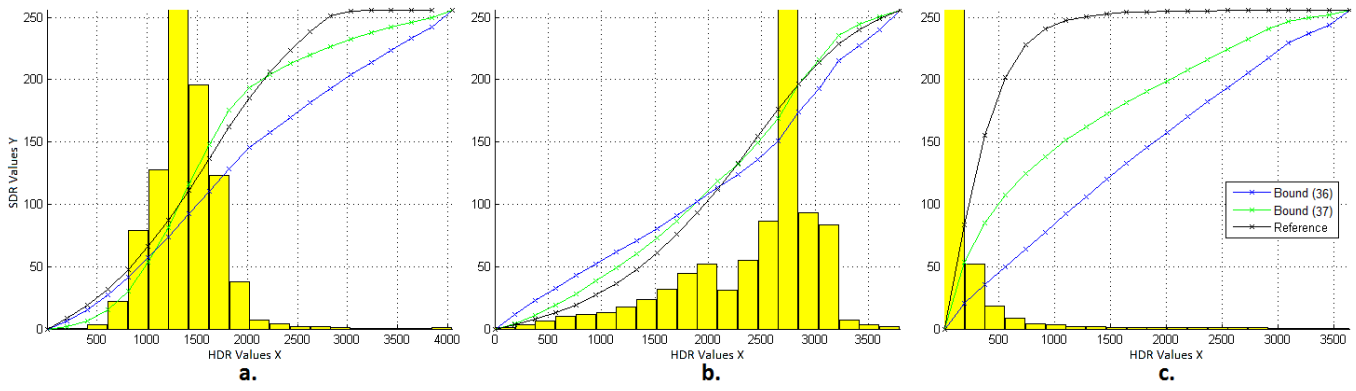


Fig. 16. Different tone mapping curves for images a. AtriumNight, b. Ballon and c. Fireater. The (36) bound by the slopes  $s_k^o$ , the (37) bound is defined by the slopes  $s_k^t$  and the SDR perceptual reference TMO is [19]

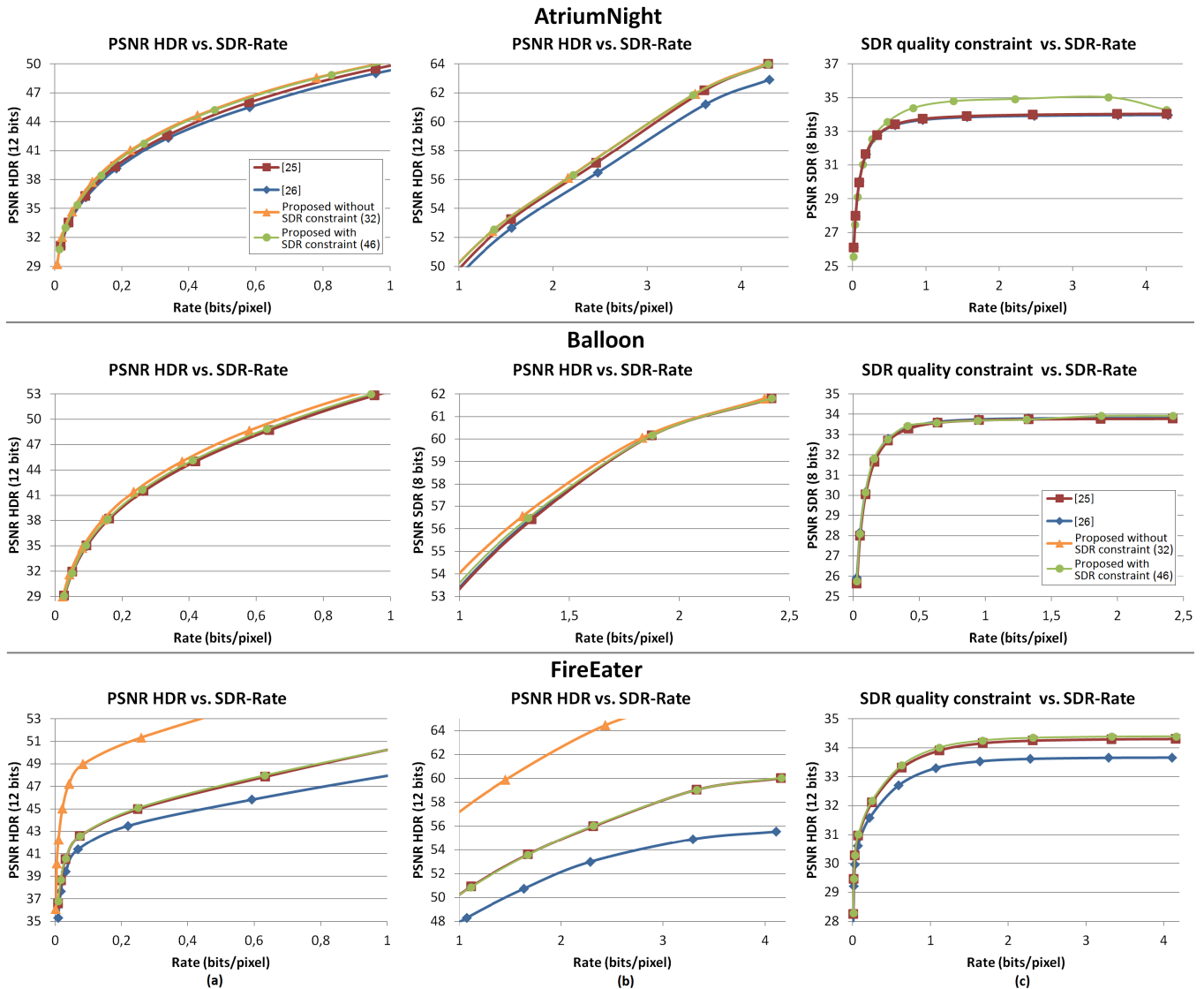


Fig. 17. Results for the images AtriumNight Ballon and FireEater. Distortion between the original HDR image  $X$  and the reconstructed image  $\tilde{X}$  vs SDR rate - (a) at low rates (b) at high rates (c) SDR quality constraint vs SDR rate

Experimental results show that the first TMO achieves the optimal rate-distortion performances with a global tone mapping and can be computed in real-time. As expected,

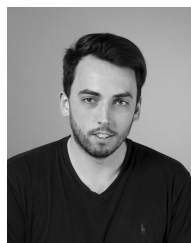
the second TMO leads to the best trade-off between rate-distortion performances and quality preservation of the SDR signal, in comparison to state of the art methods that solve

approximations of the original optimization problem. Besides, we showed that spatial gradient is an accurate estimator of both rate and distortion.

In future works, we will consider the extension to video compression addressing in particular issues of temporal consistency.

## REFERENCES

- [1] F. Banterle, A. Artusi, K. Debbastista, and A. Chalmers, "Advanced High Dynamic Range Imaging: Theory and Practice", 1st ed. Natick, MA, USA: A. K. Peters, Ltd., 2011.
- [2] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, and W. Heidrich, "High Dynamic Range Imaging - Acquisition, Display, and Image-Based Lighting (2. ed.)". Academic Press, 2010. [Online]. Available: <http://www.elsevierdirect.com/product.jsp?isbn=9780123749147&dmnum=CWS1>
- [3] G. Ward, "Real Pixels," in *Graphics Gems II*, J. Arvo, Ed. Academic Press, 1991, pp. 80–83.
- [4] I. Light and Magic, "OpenEXR, high dynamic-range (HDR) image file format," [www.openexr.com/index.html](http://www.openexr.com/index.html).
- [5] A. Motra and H. Thoma, "An Adaptive LogLuv Transform for High Dynamic Range Video Compression," in *Image Processing (ICIP), IEEE International Conference on*, 2010.
- [6] G. Larson, "LogLuv Encoding for Full-gamut High-dynamic Range Images," *J. Graph. Tools*, pp. 15–31, 1998. [Online]. Available: <http://dx.doi.org/10.1080/10867651.1998.10487485>
- [7] A. Luthra, E. Franois, and W. Husak, "Call for Evidence (CfE) for HDR and WCG Video Coding," ISO/IEC JTC1/SC29/WG11, Geneva, Switzerland, Tech. Rep. N15083, February 2015.
- [8] S. Miller, M. Nezamabadi, and S. Daly, "SMPTE 2084 - Perceptual Signal Coding for More Efficient Usage of Bit Codes," in *SMPTE Conferences*. Society of Motion Picture and Television Engineers, 2012.
- [9] T. Borer and A. Cotton, "A Display Independent High Dynamic Range Television System," in *White Paper WPH 309*. BBC British Broadcasting Corporation, 2015.
- [10] T. Borer and A. Cotton, "ARIB STD-B67 - Essential Parameter Values for the Extended Image Dynamic Range Television System for Programme Production," in *ARIB Standard*. Association of Radio Industries and Businesses, 2015.
- [11] L. Kerofsky, Y. Ye, and Y. He, "Recent developments from MPEG in HDR video compression," pp. 879–883, Sept 2016.
- [12] C. Wong, G. Su, and M. Wu, "Joint Baseband Signal Quantization and Transform Coding for High Dynamic Range Video," *CoRR Computing Research Repository*, vol. abs/1603.02980, 2016.
- [13] J. Strm, K. Andersson, M. Pettersson, P. Hermansson, J. Samuelsson, A. Segall, J. Zhao, S. H. Kim, K. Misra, A. M. Tourapis, Y. Su, and D. Singer, "High quality HDR video compression using HEVC main 10 profile," in *PCS Picture Coding Symposium*, Dec 2016, pp. 1–5.
- [14] T. Lu, F. Pu, P. Yin, Y. He, L. Kerofsky, Y. Ye, Z. Gu, and D. Baylon, "Compression Efficiency Improvement over HEVC Main 10 Profile for HDR and WCG Content," pp. 279–288, March 2016.
- [15] J. Boyce, Y. Ye, J. Chen, and A. Ramasubramanian, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20–34, Jan 2016.
- [16] T. Lu, F. Pu, P. Yin, J. Pytlarz, T. T. Chen, and W. Husak, "Adaptive resampler for high dynamic range and wide color gamut video compression," vol. 9971, 2016. [Online]. Available: <http://dx.doi.org/10.1117/12.2236050>
- [17] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of Tone Mapping Operators Using a High Dynamic Range Display," *ACM Trans. Graph.*, 2005. [Online]. Available: <http://doi.acm.org/10.1145/1073204.1073242>
- [18] Y. Salih, W. bt Md-Esa, A. Malik, and N. Saad, "Tone Mapping of HDR Images: A Review," in *Intelligent and Advanced Systems (ICIAS), International Conference on*, June 2012.
- [19] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic Tone Reproduction for Digital Images," *ACM Trans. Graph.*, 2002. [Online]. Available: <http://doi.acm.org/10.1145/566654.566575>
- [20] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, Aug. 2008.
- [21] Z. Mai, H. Mansour, R. Mantiuk, R. Ward, P. Nasiopoulos, and W. Heidrich, "Optimizing a Tone Curve for Backward Compatible High Dynamic Range Image and Video Compression," *IEEE Trans. on Image Processing*, Jun. 2011.
- [22] P. Lauga, G. Valenzise, G. Chierchia, and F. Dufaux, "Improved Tone Mapping Operator for HDR Coding Optimizing the Distortion/Spatial Complexity Trade-off," in *Signal Processing Conference (EUSIPCO), Proc. of the European*, Lisbon, Portugal, 2014. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01141081>
- [23] M. L. Pendu, C. Guillemot, and D. Thoreau, "Rate Distortion Optimized Tone Curve for High Dynamic Range Compression," in *Signal Processing Conference (EUSIPCO), Proc. of the European*, Lisbon, Portugal, 2014.
- [24] Z. Mai, H. Mansour, P. Nasiopoulos, and R. Ward, "Visually-Favorable Tone-Mapping with High Compression Performance," in *Image Processing (ICIP), IEEE International Conference on*, Sept 2010, pp. 1285–1288.
- [25] Z. Mai, H. Mansour, P. Nasiopoulos, and R. Ward, "Visually Favorable Tone-Mapping with High Compression Performance in Bit-Depth Scalable Video Coding," *Multimedia, IEEE Trans. on*, 2013.
- [26] A. Koz and F. Dufaux, "Optimized Tone Mapping with LDR Image Quality Constraint for Backward-Compatible High Dynamic Range Image and Video Coding," in *Image Processing (ICIP), IEEE International Conference on*, 2013.
- [27] A. Koz and F. Dufaux, "Methods for Improving the Tone Mapping for Backward Compatible High Dynamic Range Image and Video Coding," *Image Commun.*, vol. 29, no. 2, Feb. 2014.
- [28] D. Gommelet, A. Roumy, C. Guillemot, M. Ropert, and J. L. Tanou, "Rate-Distortion Optimization of a Tone Mapping with SDR Quality Constraint for Backward-Compatible High Dynamic Range Compression," in *Image Processing ICIP, IEEE International Conference on*, Sep 2016.
- [29] T. Cover and J. Thomas, "Elements of Information Theory", 2nd ed. Hoboken, New Jersey: Wiley-Interscience, 2006.
- [30] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.
- [31] W.-J. Kim, J.-W. Yi, and S.-D. Kim, "A Bit Allocation Method Based on Picture Activity for Still Image Coding," *IEEE Transactions on Image Processing*, vol. 8, no. 7, pp. 974–977, Jul 1999.
- [32] K. Chang, A. Men, and W. Zhang, "Adaptive Intra-frame Rate Control for H.264/AVC," in *Multimedia and Ubiquitous Engineering, Third International Conference on*, June 2009, pp. 54–58.
- [33] L. Tian, Y. Zhou, and X. Cao, "A New Rate Complexity QP Algorithm (RCQA) for HEVC Intra-Picture Rate Control," in *Computing, Networking and Communications (ICNC), International Conference on*, Feb 2014, pp. 375–380.
- [34] "Appendix and Supplementary Documents, Pictures and Results [Online]," [www.irisa.fr/temics/demos/TMOGradientBased](http://www.irisa.fr/temics/demos/TMOGradientBased).
- [35] J. Moré and D. Sorensen, "Computing a trust region step," *SIAM Journal on Scientific and Statistical Computing*, vol. 4, no. 3, pp. 553–572, 1983.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. New York, NY, USA: Cambridge University Press, 2004.
- [37] Max-Planck-Institut-Informatik, "HDR Image Gallery," [Online]. Available: [resources.mpi-inf.mpg.de/hdr/gallery.html](http://resources.mpi-inf.mpg.de/hdr/gallery.html).
- [38] G. Ward, "High dynamic range image examples," [Online]. Available: [www.anywhere.com/gward/hdrenc/pages/originals.html](http://www.anywhere.com/gward/hdrenc/pages/originals.html).
- [39] G. Valenzise, F. D. Simone, P. Lauga, and F. Dufaux, "Performance evaluation of objective quality metrics for HDR image compression," in *Applications of Digital Image Processing XXXVII*. SPIE, Aug. 2014. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01436204>



**David Gommelet** received the Engineering degree from Ecole Supérieure d'Ingénieurs de Rennes (ESIR), France, in 2012. He is currently pursuing his Ph.D. studies in Computer Science in INRIA (Institut National de Recherche en Informatique et en Automatique) and Ericsson-Envivio in Rennes, France. His current research interests include High Dynamic Range imaging and image and video compression.



**Aline Roumy** received the Engineering degree from Ecole Nationale Supérieure de l'Électronique et des Applications (ENSEA), Cergy, France in 1996, the Master degree in June 1997 and the Ph.D. degree in September 2000 from the University of Cergy-Pontoise, France. During 2000-2001, she has been the recipient of a French Defense DGA/DRET postdoctoral fellowship and was a research associate at Princeton University, Princeton, NJ. On November 2001, she joined INRIA, Rennes, France. Her current research and study interests include the area

of statistical signal processing, coding theory and information theory.



**Christine Guillemot** is currently Director of Research at INRIA (Institut National de Recherche en Informatique et Automatique) in France. She holds a PhD degree from ENST (Ecole Nationale Supérieure des Telecommunications) Paris (1992). From 1985 to 1997, she has been with France Télécom in the areas of image and video compression for multimedia and digital television. From 1990 to mid 1991, she has worked as visiting scientist at Bellcore Bell Communication research) in the USA. Her research interests are signal and image processing, and in

particular 2D and 3D image and video coding, joint source and channel coding for video transmission over the Internet and over wireless networks, and distributed source coding. She has served as Associate Editor for IEEE Trans. on Image Processing (from 2000 to 2003, and from 2014-2016), for IEEE Trans. on Circuits and Systems for Video Technology (from 2004 to 2006), and for IEEE Trans. on Signal Processing (2007-2009). She has served as senior member of the editorial board of the IEEE journal on selected topics in signal processing (2013-2015) and is currently senior area editor of IEEE Trans. on Image Processing. She is IEEE fellow since January 2013.



**Michaël Ropert** is currently heading the advanced video team at Ericsson in France. He holds a PhD degree from Rennes I University (1995). Previously, (1991-1992) he was teacher in applied mathematics. He has been with France Télécom in the areas of image and video compression for multimedia and digital television (1996-1999). He has been at the initial step of Envivio and leading the video team for several years (2000-2015).



**Julien Le Tanou** is currently a Senior engineer within Ericsson's Media Processing Business Line, conducting research into video processing and compression technologies. Since 2012, he has been with Envivio France, which was acquired by Ericsson in 2015, in charge of research and algorithms design of Envivio's real-time video encoding product, with a focus on HEVC video encoder efficiency optimization, UHD and HDR/WCG signal processing. Prior to Envivio, he has been with Orange Labs in France and Dolby Laboratories in USA, working on

new paradigms of image and video compression for contributing to future TV/video standards. He received an Engineering degree in Telecommunication and Multimedia from Institut Supérieure d'Electronique de Paris, France in 2008, and an Advanced Master's degree in Signal, Image Processing, and Pattern Recognition from Telecom ParisTech, France in 2010.